31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Nonstationary Modeling of Neural Population Dynamics

Rosa H. M. Chan, *Student Member*, *IEEE*, Dong Song, *Member*, *IEEE*, and Theodore W. Berger, *Senior Member, IEEE*

***Abstract*—A stochastic state point-process adaptive filter was used to track the temporal evolution of several simulated nonlinear dynamical systems. The estimated Laguerre coefficients and Laguerre poles were used to reconstruct the feedforward and feedback kernels in the system. Simulations showed that the proposed method could track the actual underlying changes of nonlinear kernels using spike input and spike output information alone. The estimated models also converge quickly to the actual models after abrupt step changes in kernels. The proposed method can be used to track the functional input-output properties of neural systems as a result of learning, changes in context, aging or other factors in the natural flow of behavioral events.**

I. Introduction

THE brain comprehends information and initiates actions through the ensemble spiking activity of its neurons. Biological processes underlying spike transformations across brain regions, including synaptic transmission, dendritic integration, and spike generation, are highly nonlinear dynamical processes and are often nonstationary. For example, it is well established that certain forms of synaptic plasticity, such as long-term potentiation (LTP) and long-term depression (LTD), occur in response to specific input patterns, and plasticity is manifested as a change in input-output functions that can be viewed as a system nonstationarity. Quantitative studies of how such functions of information transmission across brain regions evolve during behavior are required in order to understand the brain. This paper proposes a nonstationary modeling framework for the multiple-spike activity propagations between brain regions.

We have identified hippocampal CA3-CA1 spike train transformations in well-trained animals performing a delayed-nonmatch-to-sample (DNMS) task with multiple-input, multiple-output (MIMO) models [1]. The models we have developed can stochastically predict CA1 output spike trains based on CA3 input spike trains. In this first step toward modeling the hippocampus, the experimental data are collected from well-trained animals who achieved asymptotic performance. That is, they performed the behavioral task nearly identically each time so the underlying input-output transformations are nearly stationary. This implies that we can apply one time-invariant model to describe CA3-CA1 nonlinear dynamics. In the second step of hippocampal modeling, we tried to identify, in the learning animals, the hippocampal CA3-CA1 functional connections that are expected to change over time, and so the kernels describing the input-output dynamics also would be expected to change over time. We seek to identify such time-varying properties of hippocampal nonlinear dynamics by extending our model to be nonstationary using adaptive signal processing techniques.

In system identification applications, an adaptive filter is used to estimate a model that represents the best fit to an unknown plant. For example, the steepest descent algorithm using a deterministic gradient was applied to nonstationary systems [2]. However, optimal learning rates are related to the covariance matrix of variables, which are usually unknown, and are time-varying in nonstationary systems. Stochastic gradient algorithms [3], such as the normalized least-mean-squares (NLMS) [4] are most commonly used for nonstationary linear systems. They utilize an estimation of input power to adjust learning rates at each time step.

The standard Kalman filter, has also been applied to linear systems [4, 5]. Wu *et al.* applied the Kalman filter to control cursor motion [5]. Various modified Kalman filters have also been derived. For example, the switching Kalman filter proposed by Wu *et al.* [6] is a weighted piecewise-linear model. Inputs used are the firing rates of neurons [7].

Gao *et al.* have investigated the particle filters for brain-machine interfaces (BMI) decoding with a variety of nonlinear models [8]. However, a particle filter increases computational costs significantly. Neural Network is a popular choice [4, 9, 10] but it is sometimes hard to interpret the multiple hidden layers and nodes in networks. Marmarelis proposed the expansion of time-varying coefficients onto basis functions to characterize time-varying nonlinear systems [11]. Krieger *et al.* applied his technique to study the kernel changes in potentiation induced by rapid stimulations [12]. Chon and colleagues [13] extended this projection method to develop time-varying principal dynamic modes (PDM) [14] for the study of single-input, single output (SISO) time–varying nonlinear dynamic systems. They predetermined the number of Walsh or Legendre basis functions used. However, the optimal number of basis functions depends on the number of occurrences in parameter changes, which is unknown in time.

Another type of neural model examined the effects of

Manuscript received June 20, 2009.

R. H. M. Chan is with the Center for Neural Engineering, Department of Biomedical Engineering, University of Southern California, Los Angeles, CA 90089 USA (e-mail: homchan@usc.edu).

D. Song and T. W. Berger are with the Center for Neural Engineering, Department of Biomedical Engineering, University of Southern California, Los Angeles, CA 90089 USA

multiple external cues on individual neurons by tracking model parameters in the point process framework. Eden *et al.* proposed the stochastic state point process filter (SSPPF) [15]. The SSPPF updates model coefficients in proportion to the difference between the occurrence of an actual spike and the estimated probability of its occurrence. Single unit spike activities are regarded as the most relevant neural signals and will serve as the focus of investigation here. Therefore, SSPPF is applied to our nonlinear dynamical modeling framework to track nonstationary systems.

II. METHODS

A. Configuration of the Generalized Volterra Model (GVM)

A MIMO system can be decomposed into a series of multiple-input, single-output (MISO) systems as shown in Fig. 1 [16]. The MISO models are identical in structure and each module projects to a separate output. Each MISO model has physiologically-plausible components which can be described by the following equations:

$$w = u(k,x) + a(h,y) + \varepsilon(\sigma), \quad (1)$$

$$y = \begin{cases} 0 & \text{when } w < \theta \\ 1 & \text{when } w \geq \theta \end{cases}. \quad (2)$$

The input and output spike trains are denoted by x and y respectively. The hidden variable w represents the "pre-threshold membrane potential" of the output neuron. The "pre-threshold membrane potential" is the summation of the "synaptic potential" u, the output spike-triggered "after-potential" a, and a Gaussian white noise input ε with standard deviation σ. The noise term models both the intrinsic noise of the output neuron and contributions from unobserved inputs. When w crosses the threshold θ, an output spike is generated and a feedback after-potential a is triggered and then added to w. Consider the i^{th} order Volterra kernel k_i where N is the number of inputs and M_k is the finite memory of the kernel. Then, the "synaptic potential" u can be expressed as

$$\begin{aligned} u(t) = k_0 &+ \sum_{n=1}^{N}\sum_{\tau=0}^{M_k} k_1^{(n)}(\tau)x_n(t-\tau) \\ &+ \sum_{n=1}^{N}\sum_{\tau_1=0}^{M_k}\sum_{\tau_2=0}^{M_k} k_{2s}^{(n)}(\tau_1,\tau_2)x_n(t-\tau_1)x_n(t-\tau_2) \\ &+ \sum_{n_1=1}^{N}\sum_{n_2=1}^{n_1-1}\sum_{\tau_1=0}^{M_k}\sum_{\tau_2=0}^{M_k} k_{2x}^{(n_1,n_2)}(\tau_1,\tau_2)x_{n_1}(t-\tau_1)x_{n_2}(t-\tau_2) \\ &+ \ldots. \end{aligned} \quad (3)$$

The feedback variable a can be expressed as

$$a(t) = \sum_{\tau=1}^{M_h} h(\tau)y(t-\tau), \quad (4)$$

where h is the linear feedback kernel and M_h is the memory of the feedback process. The spike-triggered feedback captures the effects of ion channels such as calcium-activated potassium channels and also captures GABAergic feedback in the output firing patterns.

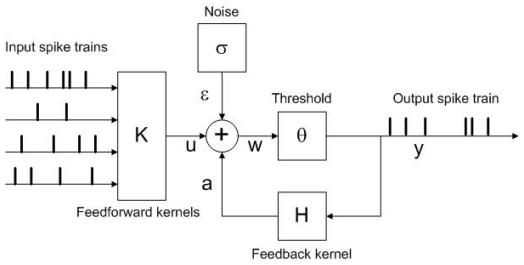


Fig. 1 MISO model structure

B. Laguerre Expansion of GVM: Generalized Laguerre-Volterra Model (GLVM)

The Laguerre expansion of the Volterra kernel (LEV) is used to reduce the number of open parameters to be estimated and, more importantly, to produce real-time prediction by separating the system's nonlinearity from the system's dynamics [16]. Using the LEV technique, both feedforward kernels k and feedback kernel h are expanded through orthonormal Laguerre basis functions b with input and output spike trains x and y convolved with b, such that the convolved products v are expressed as $v_j^{(n)} = b_j * x_n$ and $v_j^{(h)} = b_j * y$. Synaptic potential u and after potential a can be rewritten as

$$\begin{aligned} u(t) = c_0 &+ \sum_{n=1}^{N}\sum_{j=1}^{L} c_1^{(n)}(j)v_j^{(n)}(t) \\ &+ \sum_{n=1}^{N}\sum_{j_1=1}^{L}\sum_{j_2=1}^{j_1} c_{2s}^{(n)}(j_1,j_2)v_{j_1}^{(n)}(t)v_{j_2}^{(n)}(t) \\ &+ \sum_{n_1=1}^{N}\sum_{n_2=1}^{n_1-1}\sum_{j_1=0}^{L}\sum_{j_2=0}^{L} c_{2x}^{(n_1,n_2)}(j_1,j_2)v_{j_1}^{(n_1)}(t)v_{j_2}^{(n_2)}(t) \\ &+ \ldots, \end{aligned} \quad (5)$$

$$a(t) = \sum_{j=1}^{L} c_h(j)v_j^{(h)}(t). \quad (6)$$

The convolved functions v include the temporal dynamics. Another advantage of the Laguerre expansions is that the convolutions are generated in real time. Let α_n ($0 < \alpha_n < 1$) be the pole of the Laguerre basis functions for input n. The Laguerre basis functions can be obtained by inverse Z-transform of transfer function of the Laguerre filter

$$b_j^{(n)} = \mathrm{Z}^{-1}\left\{\frac{\sqrt{1-\alpha_n^2}}{1-\alpha_n z^{-1}}\left(\frac{z^{-1}-\alpha_n}{1-\alpha_n z^{-1}}\right)^{j-1}\right\}. \quad (7)$$

A Laguerre basis function of j-th order will have $j-1$ intercepts with the x-axis. The decay time of the built-in exponential of the Laguerre basis functions increases when the value of the Laguerre pole increases. v can also be computed iteratively at each time t [17]. Let $\mathbf{V}^{(n)}(t) = [v_1^{(n)}(t) \quad \ldots \quad v_L^{(n)}(t)]'$,

$$\mathbf{A_1}\mathbf{V}^{(v)}(\mathbf{t}) = \mathbf{A_2}\mathbf{V}^{(n)}(\mathbf{t-1}) + \sqrt{1-\alpha_n^2}\mathbf{A}_3 x_n(\mathbf{t}), \quad (8)$$

where $\mathbf{A}_1 = \mathbf{I} + \alpha_n \mathbf{I}_-$; $\mathbf{A_2} = \alpha_n \mathbf{I} + \mathbf{I}_-$; $\mathbf{A_3} = [1 \quad 0 \quad \ldots \quad 0]'$; $\mathbf{I}$ is an LxL identity matrix and $\mathbf{I}_-$ is a lower shift matrix.

C. Estimation of Parameters

Given recorded input and output spike trains x and y, u and a can be readily calculated based on the present values

of v and the model coefficients in real-time. The estimated firing probability $P(t)$ is calculated using the error function:

$$P(t) = 0.5 - 0.5erf\left(\frac{\theta - u(t) - a(t)}{\sqrt{2}\sigma}\right). \quad (9)$$

Without the loss of generality, θ and σ can be set to 0 and 1 respectively [1]. The parameters to be estimated are the Laguerre coefficients c and Laguerre poles α. Using the stochastic state point process filtering algorithm (SSPPF) [15], the parameter vector $C(t)$ and its covariance matrix $W(t)$ are updated iteratively at each time step t:

$$W(t)^{-1} = [W(t-1) + Q]^{-1} + \left[\left(\frac{\partial \log P(t)}{\partial C(t)}\right)' P(t)\left(\frac{\partial \log P(t)}{\partial C(t)}\right) - (y(t) - P(t))\frac{\partial^2 \log P(t)}{\partial C(t)\partial C(t)'}\right], \quad (10)$$

$$C(t) = C(t-1) + W(t)\left[\left(\frac{\partial \log P(t)}{\partial C(t)}\right)'(y(t) - P(t))\right], \quad (11)$$

where Q is a small coefficient noise covariance matrix. During adaptive parameter estimation, the gradient and the Hessian required in (10) and (11) can be also generated in real-time. The derivatives with respect to c are given as the products of v calculated in (8), such as

$$\partial u(t)/\partial c_0 = 1,$$
$$\partial u(t)/\partial c_1^{(n)}(j) = v_j^{(n)}(t),$$
$$\partial u(t)/\partial c_{2s}^{(n)}(j_1, j_2) = v_{j_1}^{(n)}(t) v_{j_2}^{(n)}(t),$$
$$\partial u(t)/\partial c_{2x}^{(n_1,n_2)}(j_1, j_2) = v_{j_1}^{(n_1)}(t) v_{j_2}^{(n_2)}(t),$$
$$\partial a(t)/\partial c_h(j) = v_j^{(h)}(t).$$

The derivatives with respect to the Laguerre poles can be calculated immediately based on the current input $x_n(t)$ and the derivatives evaluated at previous time step $t-1$. Let $\partial \mathbf{V}^{(n)}(t)/\partial \alpha_n = [\partial v_1^{(n)}(t)/\partial \alpha_n \quad \dots \quad \partial v_L^{(n)}(t)/\partial \alpha_n]'$,

$$\mathbf{A}_1 \frac{\partial \mathbf{V}^{(n)}(\mathbf{t})}{\partial \alpha_n} = \mathbf{A}_2 \frac{\partial \mathbf{V}^{(n)}(\mathbf{t}-\mathbf{1})}{\partial \alpha_n} + \mathbf{V}^{(n)}(\mathbf{t}-\mathbf{1}) - \mathbf{I}\_\mathbf{V}^{(n)}(\mathbf{t}) - \frac{\alpha_n}{\sqrt{1-\alpha_n^2}}\mathbf{A}_3 x_n(\mathbf{t}). \quad (12)$$

Let $\partial^2 \mathbf{V}^{(n)}(t)/\partial \alpha_n^2 = [\partial^2 v_1^{(n)}(t)/\partial \alpha_n^2 \quad \dots \quad \partial^2 v_L^{(n)}(t)/\partial \alpha_n^2]'$,

$$\mathbf{A}_1 \frac{\partial^2 \mathbf{V}^{(n)}(\mathbf{t})}{\partial \alpha_n^2} = \mathbf{A}_2 \frac{\partial^2 \mathbf{V}^{(n)}(\mathbf{t}-\mathbf{1})}{\partial \alpha_n^2} + 2\frac{\partial \mathbf{V}^{(n)}(\mathbf{t}-\mathbf{1})}{\partial \alpha_n} - 2\mathbf{I}\_\frac{\partial \mathbf{V}^{(n)}(\mathbf{t})}{\partial \alpha_n} - \left(1-\alpha_n^2\right)^{-\frac{3}{2}}\mathbf{A}_3 x_n(\mathbf{t}). \quad (13)$$

Since α_n is bounded between 0 and 1, we define a variable f_n as $\alpha_n = 0.5 + 0.5\, erf(f_n)$, such that parameter vector C composed of unrestrained parameters c and f_n. Other sigmoid functions can be also used to set the upper bound and the lower bound of α_n.

After observation of actual output spike train and prediction of firing probability in (9), the coefficient covariance matrix W is first estimated in (10) using the gradient and the Hessian calculated. Then W acts as the learning rate for the parameter estimations in (11). The estimated Laguerre coefficients $\tilde{c}$ and the Laguerre poles $\tilde{\alpha}$ are used to reconstruct the feedforward and feedback kernels [1].

D. Reconstruction of Kernels

The final coefficients $\hat{c}$ and $\hat{\sigma}$ can be obtained from estimated Laguerre expansion coefficients, $\tilde{c}$, through a normalization procedure by taking $\hat{c}^{(n)} = \tilde{c}^{(n)}/(\theta - \tilde{c}_0)$, $\hat{c}^{(h)} = \tilde{c}^{(h)}/(\theta - \tilde{c}_0)$, and $\hat{\sigma} = \sigma/(\theta - \tilde{c}_0)$. Feedforward kernels then can be reconstructed as:

$$\hat{k}_0 = 0,$$
$$\hat{k}_1^{(n)}(\tau) = \sum_{j=1}^{L} \hat{c}_1^{(n)}(j) b_j^{(n)}(\tau),$$
$$\hat{k}_{2s}^{(n)}(\tau_1, \tau_2) = \sum_{j_1=1}^{L}\sum_{j_2=1}^{j_1} \frac{\hat{c}_{2s}^{(n)}(j_1, j_2)}{2}\left(b_{j_1}^{(n)}(\tau_1) b_{j_2}^{(n)}(\tau_2) + b_{j_2}^{(n)}(\tau_1) b_{j_1}^{(n)}(\tau_2)\right)$$
$$\hat{k}_{2x}^{(n_1,n_2)}(\tau_1, \tau_2) = \sum_{j_1=1}^{L}\sum_{j_2=1}^{L} \hat{c}_{2x}^{(n_1,n_2)}(j_1, j_2) b_{j_1}^{(n_1)}(\tau_1) b_{j_2}^{(n_2)}(\tau_2)$$

Similar by, a linear feedback kernel can be reconstructed as

$$\hat{h}(\tau) = \sum_{j=1}^{L} \hat{c}_h(j) b_j^h(\tau).$$

The normalized kernels provide an intuitive nonparametric representation of the system input-output nonlinear dynamics. For instance, $\hat{k}_{2x}^{(n_1,n_2)}(\tau_1, \tau_2)$ represents the joint nonlinear effect of pairs of spikes with one spike from neuron n_1 and one spike from neuron n_2 [16].

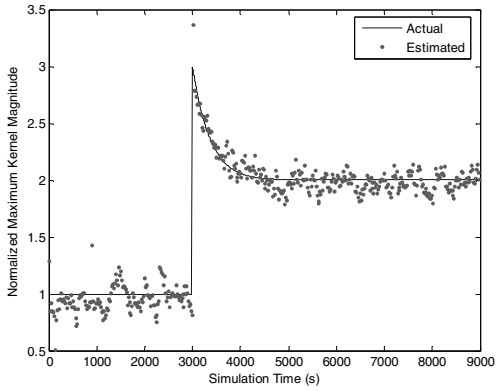
III. Results

A. Simulation of Time-Varying Neural System

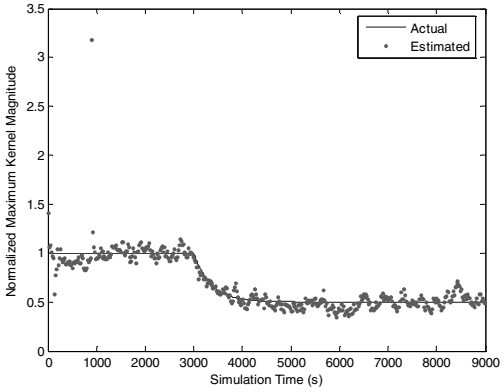
Simulations of multiple inputs were conducted [18]. By looking at the simulated point process inputs and outputs only, the changes in the underlying nonlinearities were identified. Since the Laguerre pole is also optimized, it allows us to track the kernel with unknown time window. An example of simulations is shown in Fig. 2. The system simulated was a 2-input first-order system, the kernels of the first input and the second input had LTP- and LTD-like changes respectively. After the "short-term potentiation" reached its peak magnitude, it decayed with a time constant of 6 minutes in Fig. 2a . Fig. 2b demonstrated the simultaneous tracking of a long-term depression-like time-varying kernel of the second input, which decayed with a time constant of 6 minutes.

Bin size of the spike trains was 2 ms and the kernels were tracked at every bin. The Laguerre coefficients c and Laguerre poles α_n were initialized at 0 and 0.5 respectively and they converged to the optimal values with respect to the actual input kernels during the simulation. The estimated model converged quickly to the actual system after abrupt change as shown in Fig. 2a. Other simulations also demonstrated that the nonstationary model is capable of

tracking concurrent changes in multiple kernels with unknown time windows.



(a)



(b)

Fig. 2 Peak magnitude of the first order feedforward kernel in a) LTP and b) LTD. Interval between successive estimations plotted is 20 s.

IV. Discussion

Simulations of high order systems and time-varying neural systems show that the proposed method can track the actual underlying changes of nonlinear kernels by looking at spike inputs and outputs only. In practice, adaptive modeling techniques can be applied to track the time-varying dynamics between brain subregions such as hippocampal CA3 and CA1 with the MIMO nonlinear model framework proposed here. The tracked differences in kernel functions can represent differences in the feedback or the feedforward connections, or differences in the density of voltage-dependent ionic channels, among other possibilities. Correlations between classes of kernel functions and anatomical locations within the hippocampus can be also traced. Analyses of cross-kernel evolution can indicate the developments of termination patterns in the postsynaptic dendritic tree. Large-scale simulations of a larger number of inputs and performance comparison with other existing adaptive modeling techniques will be conducted to validate this proposed nonstationary model.

References

[1] D. Song, R. H. M. Chan, V. Z. Marmarelis, R. E. Hampson, S. A. Deadwyler, and T. W. Berger, "Nonlinear dynamic modeling of spike train transformations for hippocampal-cortical prostheses," *IEEE Transactions on Biomedical Engineering*, vol. 54, pp. 1053-1066, 2007.

[2] M. T. Chian, V. Z. Marmarelis, and T. W. Berger, "Decomposition of neural systems with nonlinear feedback using stimulus-response data," *Neurocomputing*, vol. 26-7, pp. 641-654, 1999.

[3] A. H. Sayed, *Fundamentals of adaptive filtering*. New York: IEEE Press Wiley-Interscience, 2003.

[4] S. P. Kim, J. C. Sanchez, Y. N. Rao, D. Erdogmus, J. M. Carmena, M. A. Lebedev, M. A. L. Nicolelis, and J. C. Principe, "A comparison of optimal MIMO linear and nonlinear models for brain-machine interfaces," *Journal of Neural Engineering*, vol. 3, pp. 145-161, 2006.

[5] W. Wu, A. Shaikhouni, J. R. Donoghue, and M. J. Black, "Closed-loop neural control of cursor motion using a Kalman filter," presented at Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE, 2004.

[6] W. Wu, M. J. Black, D. Mumford, Y. Gao, E. Bienenstock, and J. P. Donoghue, "Modeling and decoding motor cortical activity using a switching Kalman filter," *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 933-942, 2004.

[7] W. Wu and N. G. Hatsopoulos, "Real-time decoding of nonstationary neural activity in motor cortex," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 16, pp. 213-222, 2008.

[8] Y. Gao, M. J. Black, E. Bienenstock, W. Wei, and J. P. Donoghue, "A quantitative comparison of linear and non-linear models of motor cortical activity for the encoding and decoding of arm motions," presented at Neural Engineering, 2003. Conference Proceedings. First International IEEE EMBS Conference on, 2003.

[9] N. Hatsopoulos, J. Joshi, and J. G. O'Leary, "Decoding continuous and discrete motor behaviors using motor and premotor cortical ensembles," *Journal of Neurophysiology*, vol. 92, pp. 1165-1174, 2004.

[10] M. Iatrou, T. W. Berger, and V. Z. Marmarelis, "Modeling of nonlinear nonstationary dynamic systems with a novel class of artificial neural networks," *IEEE Transactions on Neural Networks*, vol. 10, pp. 327-339, 1999.

[11] V. Z. Marmarelis, "Identification of Nonlinear Biological-Systems Using Laguerre Expansions of Kernels," *Annals of Biomedical Engineering*, vol. 21, pp. 573-589, 1993.

[12] D. Krieger, T. W. Berger, and R. J. Sclabassi, "Instantaneous Characterization of Time-Varying Nonlinear-Systems," *IEEE Transactions on Biomedical Engineering*, vol. 39, pp. 420-424, 1992.

[13] R. Zou, H. L. Wang, and K. H. Chon, "A robust time varying identification algorithm using basis functions," *Annals of Biomedical Engineering*, vol. 31, pp. 840-853, 2003.

[14] Y. Zhong, K. M. Jan, K. H. Ju, and K. H. Chon, "Representation of time-varying nonlinear systems with time-varying principal dynamic modes," *IEEE Transactions on Biomedical Engineering*, vol. 54, pp. 1983-1992, 2007.

[15] U. T. Eden, L. M. Frank, R. Barbieri, V. Solo, and E. N. Brown, "Dynamic analysis of neural encoding by point process adaptive filtering," *Neural Computation*, vol. 16, pp. 971-998, 2004.

[16] D. Song, R. H. M. Chan, V. Z. Marmarelis, T. W. Berger, R. E. Hampson, and S. A. Deadwyler, "Statistical Selection of Multiple-Input Multiple-Output Nonlinear Dynamic Models of Spike Train Transformation," presented at Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE, 2007.

[17] C. Boukis, D. P. Mandic, A. G. Constantinides, and L. C. Polymenakos, "A novel algorithm for the adaptation of the pole of Laguerre filters," *IEEE Signal Processing Letters*, vol. 13, pp. 429-432, 2006.

[18] R. H. M. Chan, D. Song, and T. W. Berger, "Tracking temporal evolution of nonlinear dynamics in hippocampus using time-varying volterra kernels," presented at Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE, 2008.